

DE 09/2154



REC'D 04 SEP 2000	
WIPO	PCT

4

**Prioritätsbescheinigung über die Einreichung
 einer Patentanmeldung**

PRIORITY DOCUMENT
 SUBMITTED OR TRANSMITTED IN
 COMPLIANCE WITH
 RULE 17.1(a) OR (b)

Aktenzeichen: 199 32 270.8

Anmeldetag: 05. Juli 1999

Anmelder/Inhaber: Friedrich-Schiller-Universität Jena, Jena/DE;
 Thomas M o o r e, Drackendorf/DE.

Bezeichnung: Verfahren zur mehrdimensionalen Analyse
 eines Proteoms

IPC: G 01 N 33/48

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 03. August 2000
Deutsches Patent- und Markenamt
Der Präsident
 Im Auftrag

Seiler

Seiler

21 (2), 121-30; Evans MJ et al.: Gene trapping and functional genomics, Trends Genet, 1997 Sep, 13 (9), 370-4]. Die Sequenzdaten sind in Datenbanken gespeichert. Die Aufklärung des Genoms eines Organismus führt letztlich "nur" zur Kenntnis des relativ statischen Informationsgehaltes des genetischen Materials für diesen Organismus. Mit den Sequenzen der cDNA- ist es prinzipiell möglich, Expressionslevel der mRNA auch zellspezifisch und umweltspezifisch zu ermitteln und damit ein Gen-expressionsmuster der RNA zu erhalten.

Aus einem Gen des Genoms können

- a) durch verschiedene Prozesse, unterschiedliche mRNA-Sorten entstehen, die für divergente Proteine kodieren, und
- b) die aus ihnen entstehenden Proteine können durch posttranslationale Modifikation eine Vielzahl außerordentlich unterschiedlich funktionierender Proteine bilden. Zu den bisher bekannten Modifikationen gehören Phosphorylierung und Dephosphorylierung, limitierte Proteolyse, Acetylierung, Methylierung, Adenylierung, Sulfatierung, Glykosylierung [McDonald LJ et al.: Enzymatic and nonenzymatic ADP-ribosylation of cysteine, Mol Cell Biochem, 1994 Sep, 138 (1-2), 221-6; Baenziger JU: Protein-specific glycosyltransferases: how and why they do it!, FASEB J, 1994 Oct, 8 (13), 1019-25; Mimnaugh EG et al.: The measurement of ubiquitin and ubiquitinated proteins, Electrophoresis, 1999 Feb, 20 (2), 418-28; Davis PJ et al.: Protein modification by thermal processing, Allergy, 1998, 53 (46 Suppl), 102-5]. Die exprimierten und modifizierten Proteine ergeben aber letztendlich das Muster, welches die Zelldifferenzierung und die Reaktion auf innere und äußere Einflüsse von Zellen beschreibt. Am augenfälligsten ist die eingeschränkte Bedeutung der Kenntnis des Genoms für die Realisierung eines definierten biologischen Zustandes, wenn man die unterschiedlichen Zellen in verschiedenen Organen und innerhalb eines

Um ein quantifizierbares "Bild" eines Proteoms zu erhalten, wird gegenwärtig folgendermaßen verfahren: In einem ersten Schritt müssen die biologischen Materialien aufgeschlossen und homogenisiert werden (mit Ausnahmen: z. B. beim Serum liegen sie in einer homogenen Lösung vor). Im zweiten Schritt erfolgt die Trennung der Proteine, im dritten die Identifizierung und im vierten die Auswertung der erhaltenen Daten [Ben RH et al.: Two dimensional electrophoresis, The state of the art and future directions, Proteome Research, New frontiers in functionel genomics, Springer 1997 Chap, 2, 13-33].

1. Aufschluß

Hierfür werden bekannte Verfahren und Anordnungen aus der Biochemie eingesetzt, wie beispielsweise Scherkrafthomogenisatoren, Ultraschallbehandlung, Hochdruckpressen. Die Schwierigkeit besteht in einem quantitativen und möglichst die Funktion der Proteine nicht zerstörenden Aufschluß, denn nur quantitativ aufgeschlossene Proteine liefern in dem nachfolgenden zweiten Schritt (Trennung und Detektion der Proteine) ein reales Bild des Probenmaterials [Rabilloud T: Solubilization of proteins in 2-D electrophoresis, An outline, Methods Mol Biol, 1999, 112, 9-19; Rabilloud T et al.: Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients, Electrophoresis, 1997 Mar-Apr, 18 (3-4), 307-16; Staudenmann W et al.: Sample handling for proteome analysis, Electrophoresis, 1998 May, 19 (6), 901-8].

2. Trennung und Detektion

Für die Trennung der Proteine des Proteoms wird gegenwärtig essentiell die zweidimensionale Gelelektrophorese verwendet. Es sind erste Versuche mit einer zweidimensionalen HPLC unternommen worden. Diese haben jedoch bisher nicht die Trennschärfe der zweidimensionalen Elektrophorese erreicht

- eingeschränkter dynamischer Bereich, der durch die Belastbarkeit der Trenngele hervorgerufen wird
- die maximal einsetzbare Proteinmenge ist auf einen Bereich von μg bis mg Protein begrenzt [James P: Of genomes and proteomes, Biochem Biophys Res Commun, 1997, Feb 3, 231 (1), 1-6]
- Einschränkung des verwendeten Probenvolumens
- die Trennung ist auf zwei Dimensionen beschränkt
- die für die Trennung benötigten Ampholyte und das Gelmaterial Acryllamid können zu Artefakten führen und dadurch zu schwer erkennbaren Fehlinterpretationen beitragen
- Proteine, die in sehr hohen Konzentrationen vorhanden sind, ergeben relativ starke Signale und überdecken solche in niedrigen Konzentrationen vorhandene, so daß eine direkte Identifikation und Quantifizierung in diesem Falle nicht möglich ist
- der Verlust der nativen Konformation im denaturierenden Trenngel bedingt den Verlust der biologisch funktionellen Eigenschaften und erschwert die Identifikation der Proteine über die Bestimmung ihrer biologischen Eigenschaften, wie beispielsweise ihrer katalytischen Aktivität oder ihrer spezifischen Bindungseigenschaften
- die Sekundäranalyse, wie die häufig eingesetzte, spezifische Proteolyse einzelner Proteine, gefolgt von Massebestimmungen, macht einen schwer automatisierbaren Extraktionschritt aus dem Gel oder von der Blotmembran notwendig.

3. Identifikation der Proteine

Hierfür werden üblicherweise die Sequenzierung, die Massenanalyse und Schätzung des isoelektrischen Punkts aus der Laufstrecke im Gel sowie die Peptidfragmentmassenanalyse nach Isolierung aus dem Gel und tryptischem Verdau in der Massenspektrometrie eingesetzt [Shevchenko A et al.: Linking

reverse-phase chromatography/electrospray tandem mass spectrometry, Protein Sci, 1998, Mar 7 (3), 706-19; Parker KC et al.: Identification of yeast proteins from two-dimensional gels: working out spot cross-contamination, Electrophoresis, 1998, Aug 19 (11), 1920-32]. Die erste Methode hat den Vorteil, daß sie einen sehr großen Massebereich bis zu 1 Mio Dalton zu analysieren erlaubt und relativ robust durchführbar ist. Allerdings kann sie nur diskontinuierlich durchgeführt werden. Die ESI-Technik hingegen kann quasi kontinuierlich an Trenntechniken angeschlossen werden und zeigt gegenwärtig einen starken Zuwachs sowohl in der Entwicklung der Applikationsbreite als auch hinsichtlich der technologischen Möglichkeiten. Die enormen Fortschritte, die in den letzten Jahren mit beiden Techniken erreicht wurden, erlauben Massenauflösungen bis zur Isotopenverteilung, also Auflösungen kleiner 1 Dalton. Damit wird ein Massenspektrum von Peptidfragmenten nach sequenzspezifischen, definiertem Proteaseverdau oder einer anderen definierten Spaltung der Proteine erhalten. Dieses Spektrum ist typisch für jedes Protein und wird zur Proteinidentifizierung in Sequenzdatenbanken von Proteinen und Expressed Sequence Tag Banken eingesetzt. Da die Identifikation des Proteins durch die präzise Identifikation der vorhergesagten Peptide nach Proteaseverdau zustande kommt, stört jede posttranslationale Modifikation der Proteine, beispielsweise durch Glykosylierung, die Erkennung. Darüber hinaus können Fragmentierungsspektren der einzelnen Peptide im Massenspektrometer Informationen über die Aminosäuresequenz der Peptide liefern. Diese Sequenzinformation kann allein oder zusammen mit den anderen bekannten Daten des Proteins zu dessen Identifizierung in einer Sequenzdatenbank genutzt werden. Dieses Verfahren zur Sequenzanalyse ist gegenwärtig auf Grund der Schwierigkeiten einer korrekten Dateninterpretation noch nicht im

Mit dem erfindungsgemäßen Verfahren ist die enge Mengenlimitation durch die Belastbarkeit der bisher verwendeten 2-D-Elektrophorese nicht mehr gegeben. Es sind Proteinmengen im Bereich einiger Gramm einsetzbar. Die Trennmatrices sind mehrfach nutzbar. Hierdurch ist eine höhere Reproduzierbarkeit der Ergebnisse erzielbar. Das eingesetzte Probenmaterial liegt in der flüssigen Phase vor und ist somit anschließenden Analyseschritten unmittelbar zugänglich. Durch den besseren Erhalt der nativen Eigenschaften während der Trennung sind analytische Verfahren, wie die Aktivitätsbestimmung, und immunologische Verfahren, die auf der nativen Konformation des Analyten beruhen, möglich. Die Trennung von Analyten mit gleichen Ladungs- und Größeneigenschaften ist in der meist verwendeten 2-D-Elektrophorese nicht möglich. Durch den Einsatz von mindestens einem weiteren Charakteristikum, wie beispielsweise der Hydrophobizität des Analyten, zur Trennung entfällt allerdings diese Einschränkung. Die Proben stehen in den Fraktionen nach der Trennung auch weiteren präparativen Arbeiten zur Verfügung.

Die Erfindung soll nachstehend anhand eines in der Zeichnung dargestellten Ausführungsbeispiels näher erläutert werden.

Es zeigen:

Fig. 1: Trennung von 1000 Proteinen in drei Dimensionen

Fig. 1a: Fraktionen 1 bis 33

Fig. 2a: Fraktionen 33/34 bis 67

Fig. 3a: Fraktionen 68 bis 100

Fig. 2: Graphische dreidimensionale Darstellung der Fraktionen gemäß Fig. 1

Fig. 1 enthält folgende tabellarische Auflistung:

Protein Nr.	Faktionen a	Faktionen b	Faktionen c
----------------	----------------	----------------	----------------

wobei in Fig. 1a die Faktionen a = 1 bis 33, in Fig. 2a die Faktionen a = 33/34 bis 67 und in Fig. 1c die Faktionen a = 68 bis 100 aufgeführt sind.

Fig. 2 zeigt ein dreidimensionales Diagramm mit den Positionen der durch Proteine besetzten Faktionen nach Fig. 1.

Patentansprüche

1. Verfahren zur mehrdimensionalen Analyse eines Proteoms, bei dem das biologische Material mit dem zu analysierenden Proteom aufgeschlossen und die zu dem Proteom gehörenden Proteine getrennt sowie quantitativ bestimmt und identifiziert werden, dadurch gekennzeichnet, daß die Proteine des Proteoms unter standardisierten Bedingungen einer Vielzahl n verschiedener Trennverfahren für $n > 2$ derart unterworfen werden, daß jeweils jede der in einem Trennschritt erhaltenen Fraktionen m_1 in einem darauf folgenden Trennschritt m_2 Fraktionen liefert, wobei nach n Trennschritten $m_1 * m_2 * \dots * m_n = M$ Fraktionen vorliegen, die mit r verschiedenen Analyseverfahren qualitativ und oder quantitativ durch an sich bekannte Identifikationsverfahren identifiziert und durch ebenfalls an sich bekannte Quantifikationsverfahren quantitativ bestimmt werden, so daß nach Zusammenfügen der Analysedaten ein n -dimensionales Abbild des Proteoms, charakterisiert durch Identifikatoren und Quantifikatoren sowie durch die Lage im n -dimensionalen Datenraum, gewonnen wird.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß als Trennverfahren Methoden, die nach der Größe der Proteine trennen, und/oder Methoden, die nach der Masse der Proteine trennen, und/oder Methoden, die nach der Ladung der Proteine trennen, und/oder Methoden, die nach der Hydrophobizität der Proteine trennen, und/oder Methoden, die nach der Form der Proteine trennen und/oder Methoden, die nach der Affinität der Proteine hinsichtlich spezifischer Liganden auch zu Antikörpern trennen, ausgewählt werden.

Belegexemplar

Darf nicht geändert werden

8. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß beim ersten Trennschritt die Fraktionen in einem definierten Raster, vorzugsweise im $n \cdot 96$ - fach Raster der Mikrotiterplattentechnologie gesammelt werden.

9. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß alle Identifikations- und Quantifikationsschritte in einem definiertem Raster, vorzugsweise dem $n \cdot 96$ - fach Raster, mit paßfähiger Liquidhandlingstechnik erfolgen.

10. Verfahren nach Anspruch 9, dadurch gekennzeichnet, daß alle Identifikations- und Quantifikationsschritte mit wenigstens vier zweidimensional angeordneten und simultan arbeitenden Pipettoren erfolgt.

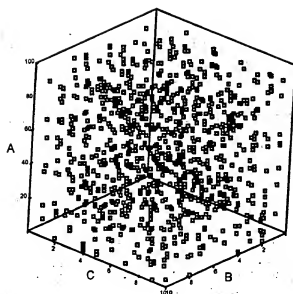
11. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die erste Dimension zur Trennung eine an sich bekannte hochauflösende Größenausschluß-, eine Ionenaustausch- oder eine Hydrophobizitätschromatographie ist, daß die zweite Dimension durch parallele Trennung und Fraktionierung der Fraktionen der ersten Dimension nach einem anderen als dem für die erste Dimension verwendeten Trennprinzip erfolgt und daß jede weitere Trennung und Fraktionierung durch parallele Trenn- und Fraktionierungsmethoden mit den aus den jeweils vorhergehenden Trenn- und Fraktionierungsschritten erhaltenen Fraktionen erfolgt.

12. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Analysedaten für das n -dimensionale Abbild des Proteoms in einer Datenbank zusammengefaßt werden.

869	33	10	3	979	41	3	5	938	48	5	10	743	55	5	7	834	62	4	2
525	34	3	3	556	41	4	9	511	48	16	54	305	55	6	8	471	62	4	3
332	34	3	4	44	41	5	3	485	48	26	54	473	55	8	1	919	62	4	10
636	34	4	6	812	41	7	8	915	48	7	5	266	55	8	4	342	62	7	3
121	34	4	8	37	41	9	6	456	48	17	58	393	55	8	10	460	62	7	8
998	34	5	1	343	41	10	8	438	48	17	58	320	55	9	1	378	62	7	10
355	34	5	8	911	42	1	6	120	48	7	10	276	55	9	6	94	62	10	7
346	34	5	10	522	42	1	10	868	48	8	3	7	55	9	10	935	63	1	7
270	34	8	1	505	42	3	1	486	48	10	10	799	55	10	9	395	63	2	4
810	34	8	2	417	42	3	2	30	49	1	6	518	56	1	10	464	63	2	8
34	34	8	1	782	42	4	3	571	49	2	1	245	56	2	6	949	63	3	2
734	34	7	1	807	42	5	7	936	49	3	10	870	56	4	5	394	63	4	1
862	34	7	6	765	42	6	3	520	49	6	10	940	56	5	3	14	63	4	6
164	34	9	4	168	42	8	1	775	49	8	4	673	56	5	6	683	63	6	4
157	34	10	8	857	42	8	9	421	49	8	6	1828	56	5	10	298	63	6	10
796	35	1	6	657	42	9	1	287	50	1	1	1628	56	5	10	598	63	7	5
962	35	1	8	252	42	10	1	89	50	3	3	679	56	8	1	8209	63	7	29
736	35	3	4	475	42	10	3	847	50	3	7	78	56	8	4	4819	63	7	29
85	35	3	8	173	42	10	4	49	50	5	8	437	56	8	8	817	63	8	8
47	35	4	2	302	43	2	3	577	50	6	7	873	56	9	6	75	63	8	9
793	35	4	6	809	43	2	10	2	50	7	5	888	56	10	5	416	63	8	10
819	35	6	6	431	43	4	3	374	50	7	7	201	57	1	7	371	63	9	3
671	35	6	8	908	43	5	10	711	50	8	9	412	57	5	9	739	63	9	5
432	35	6	10	3028	43	5	10	722	50	9	1	133	57	5	9	646	63	9	6
185	35	9	1	933	43	7	6	958	50	10	4	908	57	6	2	135	63	9	7
324	35	9	4	492	43	9	4	645	51	1	6	967	57	6	3	233	63	9	10
658	35	10	3	349	43	9	7	720	51	2	1	12	57	7	7	237	64	1	4
468	36	2	1	364	43	10	7	300	51	2	3	1672	58	1	1	764	64	4	3
643	36	4	2	197	43	10	9	973	51	6	4	139	58	1	1	974	64	1	5
926	36	5	8	465	44	2	1	282	51	6	10	352	58	1	9	622	64	1	6
693	36	8	7	549	44	2	4	674	51	7	7	293	58	2	10	745	64	6	5
767	36	9	3	635	44	3	9	213	51	9	4	543	58	9	3	248	64	6	6
354	36	10	7	538	44	5	6	833	51	9	10	954	58	10	2	585	64	6	9
955	37	1	1	801	44	6	6	216	51	10	4	681	59	2	1	466	64	9	2
314	37	4	4	993	44	6	7	986	52	1	3	844	59	2	9	217	64	9	5
548	37	5	8	965	44	7	6	253	52	2	8	753	59	3	2	730	64	9	8
313	37	6	10	780	44	8	2	625	52	2	9	881	59	5	2	761	64	10	6
219	37	7	4	830	44	9	2	768	52	3	6	52	59	5	3	569	64	10	8
959	37	8	5	277	44	9	5	818	52	3	7	501	59	5	8	750	65	1	3
46	37	9	2	269	44	9	7	804	52	6	1	516	59	6	3	38	65	2	4
497	38	1	2	113	45	1	1	824	52	6	5	196	59	6	7	102	65	3	8
678	38	3	5	265	45	1	3	705	52	6	7	860	59	7	7	880	65	4	8
260	38	3	8	710	45	1	6	512	52	6	7	628	59	8	10	528	65	5	1
754	38	3	9	477	45	2	1	337	52	8	3	162	59	10	1	725	65	7	10
648	38	4	7	922	45	2	6	639	52	9	7	54	59	10	8	787	65	8	1
310	38	4	10	668	45	4	2	204	52	9	10	291	60	2	5	533	65	9	1
209	38	4	10	271	45	4	7	284	52	10	7	553	60	2	7	408	65	9	8
990	38	6	7	863	45	6	2	615	53	1	8	655	60	2	9	892	66	1	2
11	38	8	10	39	45	7	5	261	53	1	8	2278	60	3	1	264	66	1	5
941	38	9	5	948	45	7	9	612	53	6	6	165	60	8	1	87	66	1	6
184	38	9	8	376	46	1	5	604	53	10	3	199	60	8	2	429	66	2	5
637	39	1	8	428	46	5	5	15	53	10	7	210	60	10	2	789	66	3	8
545	39	2	7	317	46	8	1	441	53	10	8	442	60	10	6	328	66	4	9
163	39	7	9	117	46	9	6	843	54	1	3	728	60	10	9	960	66	6	3
267	39	8	8	689	47	1	5	97	54	1	6	633	61	1	1	340	66	8	8
1	39	9	6	992	47	2	4	235	54	4	1	220	61	1	6	794	66	9	1
229	39	10	1	559	47	2	6	112	54	5	10	179	61	5	5	452	67	1	5
53	39	10	4	375	47	3	10	583	54	5	10	808	61	7	8	500	67	1	8
822	40	1	3	554	47	5	4	91	54	7	3	223	61	9	2	510	67	3	6
891	40	1	7	624	47	5	6	249	54	7	4	336	62	1	10	10	67	5	2
335	40	3	6	565	47	5	6	24	54	8	2	232	62	2	6	650	67	6	3
623	40	4	2	457	47	9	3	576	54	8	10	77	62	2	8	490	67	6	9
901	40	5	8	17	47	10	4	160	55	1	10	399	62	2	10	586	67	7	9
828	40	6	6	864	48	1	1	410	55	2	1	55	62	3	8	344	67	8	4
666	40	7	10	372	48	2	4	929	55	2	4	675	62	3	10	462	67	8	7
474	40	10	2	759	48	2	7	387	55	2	5								
704	40	10	3	129	48	3	4	74	55	2	9								

Fig. 1b

Belegexemplar

**Fig. 2**



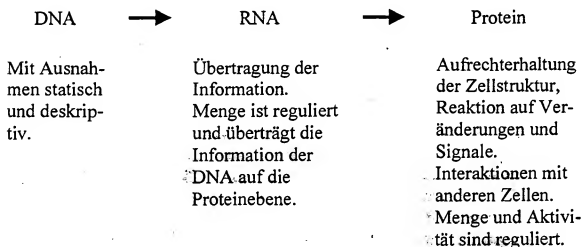
Beschreibung der Erfindung

Verfahren zur mehrdimensionalen Analyse eines Proteoms

Die Erfindung betrifft ein Verfahren zur mehrdimensionalen Analyse eines Proteoms, bei dem das biologische Gewebe mit dem zu analysierenden Proteom aufgeschlossen und die zu dem Proteom gehörenden Proteine getrennt sowie quantitativ bestimmt und identifiziert werden. Das Verfahren findet in der Biochemie, in der Biotechnologie, in der Medizin sowie in der Pharmazeutischen Industrie Verwendung und dient u. a. zu diagnostischen Zwecken und zur Entwicklung biologisch wirksamer Substanzen. Spezielle Einsatzgebiete eröffnen sich in der Grundlagenforschung, beispielsweise für die Klärung entwicklungsbiologischer oder zelldifferenzierender Fragestellungen, sowie in der angewandten Forschung für das Screening von Wirkstoffbanken, für die Entwicklung und Optimierung biologisch aktiver Substanzen oder für die Differenzierung zwischen normalen und pathogenen Zuständen in Organismen.

In der jüngeren Vergangenheit wurden Genome von Organismen ganz oder zu großen Teilen sequenziert [Fraser CM et al.: The minimal gene complement of *Mycoplasma genitalium*, Science, 1995, Oct 20, 270 (5235), 397-403; Fleischmann RD et al.: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science, 1995, Jul 28, 269 (5223), 496-512; Blattner FR et al.: The complete genome sequence of *Escherichia coli* K-12, Science, 1997, Sep 5, 277 (5331), 1453-74; Goffeau A et al.: Life with 6000 genes, Science, 1996, Oct 25, 274 (5287), 546, 563-7]. Noch intensiver wurden cDNA-Abschnitte sequenziert [Clark MS: Comparative genomics: the key to understanding the Human Genome Project, Bioessays, 1999, Feb,

Organs vergleicht. Beispielsweise haben eine Leberparanchymzelle, eine Nervenzelle des Gehirns und eine Mukosazelle des Darmes den selben Satz genetischer Information, aber völlig unterschiedliche Funktion, die durch die Regulation der Expression des Genoms in diesen Zellen und die Regulation des Enzym- und Proteinmusters innerhalb der Zellen sowie der verschiedenen Gewebe hervorgerufen wird.



Die Begriffsbestimmung des "Proteoms", erfolgte erst 1996 [Friedrich GA: Moving beyond the genome projects, Nat Biotechnol, 1996 Oct, 14 (10), 1234-7].

Das Proteom, das heißt die Gesamtheit aller Proteine in einer Zelle mit einem bestimmten Entwicklungsstand und unter definierten Umweltbedingungen, stellt eine sehr viel dynamischere Repräsentation des physiologischen Zustandes von Zellen, Organen und Organismen dar. Die Proteomanalytik untersucht, welche Teile des Genoms unter definierten, zellspezifischen Bedingungen exprimiert und modifiziert werden. Dies führte zu schnell anwachsendem Interesse an diesem Gebiet, mit der Folge von ansteigenden Publikationszahlen (PubMed query Suchbegriff: Proteome; Suche 1 Jahr zurück: 64 Einträge, 2 Jahre zurück: 99 Einträge, 5 Jahre zurück: 122 Einträge), Kongressen und Veranstaltungen zu dieser Thematik.

[Opiteck GJ et al.: Comprehensive two-dimensional high-performance liquid chromatography for the isolation of overexpressed proteins and proteome mapping., *Anal Biochem.*, 1998 May 1, 258(2), 349-61.]. Die erste Dimension der zweidimensionalen Elektrophorese ist eine Trennung nach dem isoelektrischen Punkt, also letztendlich nach den Ladungseigenschaften eines Proteins. In der zweiten Dimension wird nach der Größe der Proteine in einem denaturierenden Natriumdodecylsulfat-Gel getrennt. Diese Trenntechnik ist seit etwa seit 20 Jahren bekannt. Ein Vorteil der 2-D-Elektrophorese liegt in der Möglichkeit, eine relativ große Zahl von Proteinen auf einer Fläche mit hoher Auflösung zu trennen. Man geht gegenwärtig davon aus, daß ca. 10.000 Proteine in einem solchen zweidimensionalen Gel nachgewiesen werden können [Klose J et al.: Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome, *Electrophoresis*, 1995, Jun 16 (6), 1034-59]. Ein weiterer Vorteil liegt darin, daß man durch radioaktive Markierung oder nach der Anfärbung mit ebenfalls bekannten Techniken in der Lage ist, die getrennten Proteine zu quantifizieren. Diese Quantifizierungsmethoden sind proteinspezifisch, haben einen eingeschränkten dynamischen Nachweisbereich, sind in der Regel schwer automatisierbar und sind abhängig von den jeweiligen (oft nicht vollständig zu reproduzierenden) Einsatzbedingungen [James P: Of genomes and proteomes, *Biochem Biophys Res Commun*, 1997, Feb 3, 231 (1), 1-6]. Sie sind nur für relative Bestimmungen geeignet. Die Quantifizierung über immunologische Eigenschaften ist problematisch, weil dafür Blottechniken mit eingeschränkter quantitativer Aussagekraft eingesetzt werden müssen.

Das Ergebnis ist ein fingerabdruckähnliches Bild, welches das Proteom charakterisiert.

Die Nachteile dieser Trenntechnik sind:

genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels, Proc Natl Acad Sci USA, 1996, Dec 10, 93 (25), 14440-5; Traini M et al.: Towards an automated approach for protein identification in proteome projects, Electrophoresis, 1998, Aug 19 (11), 1941-9]. Durch die verwendete Trenntechnik werden Merkmale, wie beispielsweise die katalytische Aktivität der Proteine und die native Konformation, nahezu vollständig ausgeschaltet und stehen nicht für die Identifikation zur Verfügung.

Die Vor- und Nachteile der bekannten Identifikationsverfahren sind insbesondere:

- Die Sequenzierung erfolgt durch Edman Abbau an automatisierten Einrichtungen und ist relativ kosten- und zeitaufwendig. Sie erfordert größere Mengen des Proteins. Deshalb ist sie trotz gegenwärtiger Weiterentwicklung für ein Massenscreening weniger geeignet [Gooley AA et al.: A role for Edman degradation in proteome studies, Electrophoresis, 1997, Jun 18(7), 1068-72]. Für die Identifizierung von primär unbekannten Proteinen ist dieser Analyseschritt allerdings in den meisten Fällen notwendig.
- Die Spezifität der Aussage der Massenbestimmung eines Proteins, die letztlich zu seiner Identifizierung führen soll, wird dadurch erhöht, daß man die Proteine nach der Trennung einem Proteaseverdau unterwirft, und die mittels Masseanalytik erhaltenen Informationen mit den aus der Primärstruktur vorhergesagten Massen der Peptidsequenzen nach dem tryptischen Verdau vergleicht. Im wesentlichen werden zwei Arten der Massenspektrometrie eingesetzt. Das sind erstens die Matrix assistierte Laserdesorption/ionisierung - Massenspektrometrie (MALDI-MS) und zweitens die Electrospray Ionisierung - Massenspektrometrie (ESI-MS) [Ducret A et al.: High throughput protein characterization by automated

Routineeinsatz. Die Grenzen der Proteinidentifizierung durch massenspektrometrische Methoden bestehen in der nicht vollständigen Erfassung aller Proteinsequenzen in den vorhandenen Datenbanken.

4. Datenanalyse

Die erhaltenen Charakteristika der einzelnen detektierten Proteine aus der Trennung in der 2-D- Elektrophorese, wie die Quantität, isoelektrischer Punkt und Größe, die Daten zur Proteinidentifizierung aus weiteren Schritten, beispielsweise der Sequenzierung oder Massenspektrometrie, werden zusammengeführt. Hieraus ergibt sich das Bild der Gesamtheit der Proteine mit ihrer Identität und Quantität in dem jeweiligen Proteom.

Der Erfindung liegt die Aufgabe zu Grunde, die Quantifikation und Identifikation der Proteine eines Proteoms zu verbessern, zu erleichtern und für bestimmte Proteine überhaupt erst zu ermöglichen.

Erfindungsgemäß werden die Proteine des Proteoms unter standardisierten Bedingungen einer Vielzahl n verschiedener Trennverfahren derart unterworfen, daß jeweils jede der in einem Trennschritt erhaltenen m_1 Fraktionen in einem darauf folgenden Trennschritt m_2 Fraktionen liefert, wobei nach n Trennschritten $m_1 * m_2 * \dots * m_n = M$ Fraktionen vorliegen, die mit r verschiedenen Analyseverfahren qualitativ und oder quantitativ durch an sich bekannte Identifikationsverfahren identifiziert und durch ebenfalls an sich bekannte Quantifikationsverfahren quantitativ bestimmt werden, so daß nach Zusammenfügen der Analysedaten ein n -dimensionales Abbild des Proteoms, charakterisiert durch Identifikatoren und Quantifikatoren sowie durch die Lage im n -dimensionalen Datenraum, gewonnen wird.

In den Unteransprüchen 2-12 sind vorteilhafte Ausführungsformen des Verfahrens aufgeführt.

14

Als Ausführungsbeispiel sollen 1000 Proteine durch drei Eigenschaften A, B, C beschrieben werden. Diese Eigenschaften können zum Beispiel Größe, Ladung und Hydrophobizität sein. Die Eigenschaften sind in den Proteinen zufällig verteilt. Alle Proteine sind fortlaufend numeriert. Hierauf erfolgt eine Trennung nach der Eigenschaft A (beispielsweise der Größe), bei der 100 Fraktionen a mit den entsprechenden Proteinen erhalten werden. Diese Fraktionen a werden nach der Eigenschaft B (beispielsweise der Ladung) in jeweils 10 Fraktionen b getrennt.

Jede dieser Fraktionen b wird einer Trennung nach der Eigenschaft C (beispielsweise der Hydrophobizität) unterworfen und liefert die Fraktionen c. Insgesamt werden $100 \times 10 \times 10 = 10.000$ einzelne Fraktionen erhalten. Jedes durch die Trennung erhaltene Protein wird nach seinen Eigenschaften eindeutig einer der Fraktionen a, b, c zugeordnet. In der Aufstellung gemäß Fig. 1 sind die jeweiligen Fraktionen durch Zahlen bezeichnet. Hierbei sind die Fraktionen a der Eigenschaft A zugehörig. Sie teilen den möglichen Wertebereich der Eigenschaft A in jeweils einhundert gleiche Teile, d. h. für die Voraussetzung eines Wertebereichs von 0 bis 100 entspricht beispielsweise der Wert 1 dem Bereich 0 bis 1, der Wert 2 dem Bereich 1 bis 2, ..., der Wert 100 dem Bereich 99-100. Analog sind die möglichen Wertebereiche der Eigenschaften B und C in jeweils zehn gleiche Teile eingeteilt, d. h. beispielsweise der Wert 1 entspricht dem Bereich 1-10. Durchschnittlich jede zehnte Fraktion enthält ein Protein.

Aus der Zufallsbetrachtung ergibt sich die Möglichkeit von Mehrfachbesetzungen. In dem in der Aufstellung nach Fig. 1a-c aufgeführten Beispiel sind 39 Doppelbelegungen und eine Dreifachbelegung von Fraktionen enthalten.

Aus Platzgründen und der Übersicht halber sind die leeren 9.000 Fraktionen nicht dargestellt.

Aufstellung der verwendeten Bezugszeichen

A, B, C	-	Eigenschaft von Proteinen
a, b, c	-	Fraktion

3. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß als Identifikationsverfahren Methoden zur Bestimmung spezifischer immunologischer Eigenschaften und/oder Bestimmungsmethoden spezifischer katalytischer Aktivität und/oder Bestimmungsmethoden zur chemischen Modifikation der Proteine des Proteoms eingesetzt werden.

4. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß als Quantifikationsverfahren Methoden der unspezifischen Bestimmung der Proteinkonzentration mit unterschiedlichen Empfindlichkeiten und/oder quantitative Bestimmungsmethoden zur Bestimmung spezifischer katalytischer Aktivitäten und/oder quantitative immunologische Methoden und/oder quantitative Bindungsassays ausgewählt werden.

5. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Identifikation einzelner Proteine des Proteoms direkt über die Massebestimmung der Proteine erfolgt.

6. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Identifikation einzelner Proteine des Proteoms nach Protease Verdau und Masseidentifikation der Fragmente erfolgt.

7. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß beim ersten Trennschritt die Fraktionen in einem zweidimensionalen Mehrfachgefäßsystem, vorzugsweise in der Art und mit der Grundfläche von Mikrotitrationsplatten gesammelt werden.

Prot.	Frak	Fra	Fra	373	6	10	4	236	14	3	3	634	220	8	10	647	28	2	10
Nr.	a	k.b	k.c	699	7	1	1	724	14	4	7	640	220	8	10	242	28	4	8
404	1	1	9	508	7	1	8	151	14	6	7	385	20	8	5	757	28	5	7
218	1	3	3	480	7	1	9	275	14	7	1	316	20	10	10	450	28	6	3
970	1	4	1	582	7	2	10	304	14	7	6	70	21	1	3	707	28	9	2
777	1	4	10	531	7	2	10	726	14	7	10	470	21	2	5	611	28	9	10
325	1	5	1	177	7	6	9	664	14	10	6	581	21	2	5	428	28	9	10
190	1	5	7	918	7	7	4	214	15	1	7	604	21	3	3	463	29	1	6
348	1	6	2	239	8	1	6	525	15	1	7	698	21	3	3	259	29	2	4
459	1	6	4	649	8	1	9	263	15	2	4	795	21	4	4	19	29	4	3
156	1	8	6	106	8	2	1	181	15	2	10	23	21	6	5	530	29	4	5
172	1	8	7	610	8	2	6	289	15	3	5	811	21	8	8	722	29	4	5
999	1	8	10	738	8	5	3	731	15	5	3	29	21	9	5	60	29	5	1
588	1	9	9	154	8	5	8	547	15	6	3	362	21	10	3	603	29	5	2
189	1	9	9	154	8	7	5	877	15	6	10	709	22	1	6	907	29	6	5
281	2	1	1	981	8	8	5	631	15	7	6	347	22	2	8	660	29	7	4
618	2	2	10	972	9	3	5	596	15	7	7	695	22	3	1	685	29	8	7
274	2	4	10	96	9	3	7	202	15	7	10	663	22	4	4	411	29	10	4
849	2	5	2	192	9	4	4	535	15	9	5	597	22	4	7	829	29	10	5
494	2	6	2	694	9	6	12	193	15	10	5	579	23	2	6	797	30	3	3
81	2	7	3	691	9	6	12	800	16	1	7	614	23	2	10	383	30	3	6
564	2	10	2	952	9	6	8	187	16	1	8	434	23	3	9	600	30	3	8
670	2	10	4	68	9	8	1	560	16	3	3	22	23	8	3	472	30	4	9
562	3	1	8	913	9	8	9	98	16	5	4	420	23	8	5	32	30	5	2
805	3	2	3	950	9	9	5	714	16	6	1	983	23	10	5	542	30	6	3
108	3	2	9	150	9	10	6	680	16	7	3	975	23	10	6	953	30	6	6
923	3	3	3	137	10	1	9	652	16	8	7	244	24	1	1	391	30	8	9
386	3	3	5	80	10	2	1	454	16	8	8	688	24	2	2	171	30	10	4
295	3	4	5	969	10	4	1	390	16	8	8	561	24	3	9	123	30	10	8
36	3	7	1	778	10	5	3	876	16	9	5	257	24	4	4	224	31	1	2
20	3	10	4	462	10	7	8	369	16	9	8	5	24	5	5	418	31	2	10
606	3	10	5	737	10	8	8	116	16	10	2	280	24	6	8	964	31	2	10
592	4	1	3	785	10	9	9	770	17	2	4	212	24	8	7	773	31	3	7
8	4	2	6	802	10	10	3	902	17	3	6	896	25	2	2	762	31	4	2
917	4	2	8	856	11	1	1	988	17	3	7	389	25	2	5	159	31	5	3
756	4	3	8	467	11	6	8	957	17	4	4	815	25	3	8	73	31	5	8
255	4	4	8	932	11	9	4	183	17	4	10	331	25	4	5	409	31	7	9
483	4	4	9	927	11	10	1	946	17	5	8	589	25	6	8	783	31	8	7
503	4	6	2	326	11	10	8	402	17	5	8	814	25	6	9	370	31	9	5
991	4	6	9	594	11	10	9	823	17	5	7	816	25	7	2	874	31	10	4
741	4	9	8	523	12	2	7	593	17	9	1	87	25	7	4	733	32	1	10
507	4	10	7	116	12	3	10	977	18	3	8	845	25	7	10	100	32	2	4
69	5	1	4	297	12	4	5	221	18	6	2	890	25	9	5	502	32	2	9
584	5	2	4	447	12	4	7	498	18	6	4	104	25	9	8	381	32	3	1
26	5	2	10	238	12	8	3	752	18	8	6	491	25	10	4	786	32	3	10
771	5	3	4	605	12	8	5	272	18	8	6	568	26	1	8	638	32	4	5
515	5	3	10	558	12	8	8	996	18	8	3	379	26	2	6	859	32	4	6
188	5	5	6	169	12	9	5	438	18	8	7	909	26	3	8	599	32	5	8
566	5	6	5	43	12	9	5	521	18	9	9	687	26	4	8	448	32	5	8
478	5	7	9	987	12	10	3	968	18	10	2	105	26	5	1	321	32	5	8
529	5	8	3	191	12	10	9	735	18	10	10	858	26	6	5	871	32	7	7
149	5	8	3	866	13	1	6	90	19	1	4	827	26	6	8	118	32	9	6
627	5	8	7	997	13	2	6	443	19	1	7	798	26	7	10	622	32	9	10
729	5	9	4	852	13	2	9	361	19	2	10	769	27	1	5	607	32	10	7
956	5	9	8	142	13	3	5	3	19	3	8	976	27	3	1	84	33	1	3
832	5	10	4	613	13	4	6	436	19	3	10	415	27	3	6	83	33	3	1
897	5	10	5	338	13	5	4	662	19	6	4	296	27	3	10	939	33	3	2
792	5	10	10	890	13	5	16	79	19	8	3	921	27	4	1	925	33	4	7
350	6	1	6	43	13	5	16	273	19	9	6	99	27	4	6	16	33	5	3
360	6	3	2	669	13	6	1	403	19	9	8	851	27	4	8	653	33	5	4
423	6	3	9	903	13	7	1	493	19	10	1	240	27	5	9	208	33	6	3
875	6	4	1	656	13	7	8	524	19	10	2	519	27	8	2	567	33	7	2
82	6	5	2	537	13	8	10	914	19	10	4	766	27	8	5	194	33	8	1
578	6	5	8	779	13	9	8	246	20	1	1	51	27	8	7	58	33	8	2
487	6	7	6	708	14	1	6	272	20	3	1	451	27	8	9	499	33	8	10
363	6	8	1	141	14	2	8	985	20	6	2	198	27	8	10	198	33	8	10
333	6	9	9					841	20	6	3	288	27	9	10	700	33	9	3

Fig. 1a

Ralanavamlar

651	66	1	4
368	68	3	8
703	68	7	5
228	68	7	5
333	68	7	5
706	68	7	5
158	68	7	8
178	68	9	3
788	68	9	10
145	69	1	4
250	69	3	10
131	69	5	6
425	69	6	8
854	69	8	3
303	69	8	5
899	69	9	9
186	69	10	1
721	69	10	4
455	70	1	2
125	70	1	7
122	70	2	6
256	70	4	1
928	70	4	2
842	70	4	4
484	70	4	5
308	70	4	8
222	70	5	8
641	70	6	3
740	70	6	4
56	70	7	3
620	70	7	9
101	70	8	2
865	70	8	3
648	70	8	8
595	70	9	2
898	70	9	9
716	70	9	10
329	70	10	9
885	71	2	8
18	71	2	9
676	71	3	3
702	71	3	8
128	71	7	4
140	71	7	6
13	71	8	2
621	71	8	3
642	71	8	5
334	71	9	4
489	71	10	2
742	72	3	1
632	72	3	4
247	72	3	8
506	72	4	3
732	72	4	9
853	72	8	3
392	72	9	3
92	72	9	5
552	72	2	3
723	73	4	3
994	73	4	4
945	73	6	8
665	73	6	8
414	73	7	6
879	73	10	2
661	74	2	2
984	74	3	2
905	74	5	1
887	74	5	8
461	74	6	7
546	74	6	8
575	74	7	9
45	74	7	10
290	75	1	4
846	75	2	1
440	75	6	7
155	75	7	10
31	75	8	5
9	75	9	8
127	75	10	4
748	76	1	8
405	76	2	9
555	76	3	6
180	76	5	8
130	76	10	3
312	76	10	4
718	77	1	9
889	77	2	8
722	77	2	8
444	77	2	8
380	77	4	9
64	77	5	10
982	77	8	5
825	77	10	3
840	78	2	1
306	78	2	5
865	78	6	4
170	78	7	3
839	78	8	7
59	79	1	3
644	79	4	4
307	79	4	9
327	79	5	5
251	79	5	6
544	79	6	2
230	79	6	4
341	79	7	7
763	79	8	2
132	79	9	1
629	79	9	6
406	80	1	2
850	80	3	1
279	80	3	5
884	80	5	8
572	80	5	10
947	80	6	3
427	80	10	9
448	81	1	3
616	81	1	7
175	81	2	2
654	81	2	4
205	81	4	8
206	81	4	10
872	81	7	2
93	81	7	6
791	81	8	4
469	81	8	5
262	82	2	7
514	82	3	1
776	82	3	8
377	82	3	10
203	82	4	7
148	82	5	10
1000	82	6	5
900	82	6	8
413	82	6	10
110	82	8	5
551	83	1	2
439	83	1	9
445	83	3	1
353	83	3	2
294	83	3	10
883	83	3	10
580	83	7	2
715	83	8	3
4	83	8	4
831	83	8	6
590	83	10	5
323	83	10	8
63	84	5	2
989	84	6	5
152	84	8	9
207	84	9	2
806	84	10	2
111	84	10	7
513	85	1	2
527	85	1	10
109	85	2	7
95	85	3	8
315	85	3	9
130	85	4	6
241	85	5	5
601	85	7	3
504	85	8	4
944	85	10	9
755	86	1	1
534	86	1	5
447	86	10	2
448	86	10	2
751	86	3	5
667	86	3	8
46	86	5	2
258	86	5	8
185	86	6	8
358	86	8	2
820	86	8	4
254	86	9	1
971	86	9	4
496	86	9	8
701	86	10	7
659	87	1	5
943	87	1	8
894	87	2	5
517	87	3	3
176	87	5	5
322	87	5	10
878	87	6	4
243	87	7	8
821	87	7	9
836	87	8	2
942	87	9	6
174	87	10	2
937	88	1	3
626	88	5	1
684	88	6	8
609	88	7	2
826	88	9	5
88	89	2	1
330	89	3	5
790	89	4	1
388	89	4	8
588	89	6	5
228	89	6	6
838	89	7	5
456	89	8	2
435	89	8	7
479	89	10	5
980	90	1	2
424	90	1	3
837	90	4	5
143	90	4	10
57	90	5	2
144	90	6	6
430	90	7	7
319	90	8	2
400	90	8	4
951	90	8	7
719	90	9	10
345	90	10	3
366	90	10	8
749	90	10	10
339	90	10	10
165	90	10	10
831	91	2	2
278	91	2	7
136	91	2	10
453	91	3	2
407	91	8	3
682	91	8	4
886	91	8	7
781	91	9	5
672	91	10	4
384	92	1	2
855	92	2	5
75	92	4	3
285	92	4	10
169	92	5	1
356	92	5	9
6	92	6	2
299	92	6	6
185	92	6	9
758	92	7	4
495	92	10	2
912	92	10	6
359	93	2	6
747	93	2	9
40	93	5	2
301	93	5	3
433	93	5	7
861	93	6	6
234	93	7	1
541	93	8	1
200	93	10	1
587	93	10	8
1619	94	2	1
149	94	2	1
134	94	1	9
357	94	1	10
933	94	2	2
617	94	2	3
963	94	2	10
103	94	3	4
41	94	3	7
25	94	8	3
608	94	9	4
1634	95	2	7
1618	95	2	7
550	95	2	4
574	95	3	5
27	95	6	8
717	95	6	7
71	95	9	2
995	95	9	3
744	95	9	4
924	95	9	8
583	96	1	5
895	96	4	7
50	96	5	7
813	96	5	9
916	96	7	3
893	96	7	4
686	96	7	8
365	96	7	10
231	96	8	5
21	96	9	4
713	96	9	6
367	96	10	4
351	97	2	9
774	97	3	4
476	97	5	1
422	97	5	10
362	97	7	1
124	97	8	2
891	97	8	5
697	97	9	5
225	97	10	3
112	97	10	5
760	98	2	1
536	98	2	8
488	98	3	7
692	98	5	1
401	98	5	2
591	98	5	4
153	98	5	10
292	98	7	1
286	98	7	2
396	98	7	4
509	98	8	8
397	98	9	6
557	98	10	4
978	99	1	3
803	99	4	7
892	99	4	8
746	99	5	7
215	99	7	1
28	99	7	4
930	99	8	5
340	99	9	4
835	99	10	8
925	99	10	8
916	100	2	6
269	100	2	6
211	100	2	10
910	100	4	2
309	100	4	6
419	100	5	2
784	100	8	1
42	100	8	6
115	100	8	8
966	100	9	8
573	100	9	9
532	100	10	5

Fig. 1c

Belegexemplar